

VŠB – Technická univerzita Ostrava
Fakulta elektrotechniky a informatiky
Katedra informatiky

Sledování zdravotního stavu řidiče pomocí obrazů

Monitoring driver's health condition using images

Zadání bakalářské práce

Student:

Jan Pyszko

Studijní program:

B2647 Informační a komunikační technologie

Studijní obor:

2612R025 Informatika a výpočetní technika

Téma:

**Sledování zdravotního stavu řidiče pomocí obrazů
Monitoring of Driver's Health Condition Using Images**

Jazyk vypracování:

čeština

Zásady pro vypracování:

Sledování zdravotního stavu řidiče za pomoci obrazů může být využito zejména v automobilech, které jsou vybaveny kamerami. Například pro zvýšení bezpečnosti v silničním provozu, při monitorování únavy a zdravotního stavu řidiče a k vyhodnocování jeho schopnosti řídit automobil.

1. Seznamte se se základními pojmy v oblasti detekce a rozpoznání objektů v obrazech.
2. Seznamte se zejména s metodami pro detekci tváří a jejich částí v obrazech.
3. Implementujte detektor, který bude vyhodnocovat stav řidiče (například náhlé změny ve výrazu tváře).
4. Využít můžete například knihovny OpenCV nebo Dlib.
5. Experimentálně ověřte funkčnost, přesnost a rychlost detektoru.
6. Své závěry řádně zdokumentujte v textu práce.

Seznam doporučené odborné literatury:

- [1] A.T. Lopes, E. de Aguiar, A.F. De Souza, T. Oliveira-Santos: Facial expression recognition with convolutional neural networks: coping with few data and the training sample order, Pattern Recognit., 61, pp. 610-628 (2017)
- [2] Tie, Y., Guan, L.: Automatic landmark point detection and tracking for human facial expressions. EURASIP J. Image Video Process. 2013(1), 8 (2013)
- [3] Frigieri E., Borghi G., Vezzani R., Cucchiara R.: Fast and Accurate Facial Landmark Localization in Depth Images for In-Car Applications. In: Battiato S., Gallo G., Schettini R., Stanco F. (eds) Image Analysis and Processing - ICIAP 2017. Lecture Notes in Computer Science, vol 10484. Springer (2017)

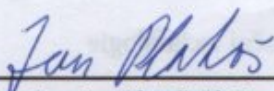
Formální náležitosti a rozsah bakalářské práce stanoví pokyny pro vypracování zveřejněné na webových stránkách fakulty.

Vedoucí bakalářské práce: **Ing. Radovan Fusek, Ph.D.**

Datum zadání: 01.09.2018

Datum odevzdání: 30.04.2019





doc. Ing. Jan Platoš, Ph.D.
vedoucí katedry



prof. Ing. Pavel Brandštetter, CSc.
děkan fakulty

Prohlašuji, že jsem tuto diplomovou práci vypracoval samostatně. Uvedl jsem všechny literární
prameny a publikace, ze kterých jsem čerpal.

V Ostravě 30. dubna 2019

.....*hond*.....

Souhlasím se zveřejněním této diplomové práce dle požadavků čl. 26, odst. 9 Studijního a zkušebního řádu pro studium v magisterských programech VŠB-TU Ostrava.

V Ostravě 30. dubna 2019

.....*Handwritten signature*.....

Rád bych poděkoval svému vedoucímu Ing. Radovanu Fuskovi za cenné rady, které mi pomohly při vytváření této práce.

Abstrakt

Tato bakalářská práce se zabývá sledováním zdravotního stavu řidiče. Implementovaný software se zabývá monitorováním a analýzou dat v oblasti tváře. Je navržený algoritmus, který získává orientační body tváře řidiče, podle kterých se určuje aktuální zdravotní stav řidiče. Ten využívá příznaky HOG a model zájmových bodů tváře, který je vytvořený pomocí datové sady 300-W. Pro sledování zdravotního stavu řidiče je i implementovaná metoda strojového učení SVM. Aby byla otestovaná funkčnost a úspěšnost, byla vytvořena aplikace v jazyce C++ s využitím knihovnem OpenCV a DLib.

Klíčová slova: SVM, Dlib, OpenCV, HOG, 300-W

Abstract

This bachelor thesis is about tracking driver's health care. Implemented software is focused on monitoring and analyzing data of face area. Algorithm is designed so it obtains driver's facial landmarks and its used to get driver's health condition. Algorithm uses HOG and shape predictor trained on dataset 300-W. For one of the symptoms is also used Support Vector Machine. To test functionality and succes rate I created application i C++ language using libraries such as OpenCV and Dlib.

Key Words: SVM, Dlib, HOG, OpenCV, HOG, 300-W

Obsah

Seznam použitých zkratk a symbolů	9
Seznam obrázků	10
Seznam tabulek	11
1 Úvod	12
2 Detekce tváře	13
2.1 Haarovy příznaky	14
2.2 Histogram orientovaných gradientů (Histogram oriented gradients)	15
2.3 Local Binary Pattern	17
3 Zájmové body tváře	18
3.1 Regrese lokálních binárních znaků	18
3.2 One Millisecond Face Alignment	19
3.3 Aktivní model tvarů (ASM)	19
3.4 Datové sety	20
4 Klasifikace akcí	23
4.1 SVM	24
5 Vlastní implementace a testování	25
5.1 Použité knihovny	25
5.2 Postup práce	26
5.3 frekvence mrkání, zívání a mikrosprávek	29
5.4 klasifikace asymetrie	30
5.5 Třes	31
5.6 Naklonění	31
5.7 Testování	32
5.8 Shrnutí testování	34
6 Závěr	36
Literatura	37
Přílohy	40
A Příloha v IS EDISON	40

Seznam použitých zkratek a symbolů

HOG	– Histogram of oriented gradients
LBF	– Local binary features
CNN	– Convolutional neural network
SVM	– Support vector machine
ASM	– Active Shape Model
3DFAW	– 3D Face Aligment in Wild
LFPW	– Labeled Face Parts in the Wild
AFLW	– Annotated Facial Landmarks in the Wild
COFW	– Caltech Occluded Faces in the Wild
300-W	– 300 Faces In-the-Wild

Seznam obrázků

1	Frontální snímání tváře řidiče. [1]	12
2	Blokové schéma. [3]	13
3	Typy příznaků (převzato z [5])	14
4	První dva příznaky vybrány AdaBoostem Převzato z: [5]	14
5	Blokové schéma HOG (převzato z [6])	15
6	Zobrazení gradientů každého pixelu jedné buňky (převzato z [8])	15
7	Přiřazování gradientů do histogramu (převzato z [8])	16
8	Převod blok pixelů do dekadické hodnoty (převzato z [9])	17
9	Ukládání počet stejných výsledků do gradientu (převzato z [9])	17
10	Dlib zájmové body (převzato z [11])	18
11	Vylepšování bodů kaskádově (převzato z [12])	18
12	Zlepšování odhadu na základě regresních stromů (převzato z [13])	19
13	Ukázka ASM (převzato z [15])	19
14	Ukázka obrázku z 300-W sady [16]	20
15	Ukázka obrázků z HELEN sady [17]	20
16	Ukázka obrázku z 3DFAW sady [18]	21
17	Ukázka anotace LFPW. Převzato z: [19]	21
18	Ukázka anotace AFLW. Převzato z: [20]	22
19	Ukázka COFW. Převzato z: [21]	22
20	Vzdálenosti bodů od těžiště [22]	23
21	Blokové schéma klasifikace. Inspirováno z: [23]	23
22	Ukázka lineárně a nelineárně rozdělitelných dat	24
23	Blokové schéma implementace.	26
24	Získaná tvář ohraničená v červeném rámci.	27
25	Reprezentace 68 zájmových bodů.	28
26	Různé pohledy a viditelné změny bodů	28
26	Ukázka zívnutí, mrknutí, mirkospánek	30
27	Asymetrie první metody	30
28	Asymetrie druhé metody	31
29	Naklonění hlavy	32
29	Chybně určené zájmové body	35

Seznam tabulek

1	Výsledky testování nesymetrie tváře obou metod	32
2	Výsledky falešně pozitivních asymetrií	32
3	Výsledky detekce mrkání	33
4	Výsledky časové úspěšnosti detekce mikrosněpánku	33
5	Výsledky zívnutí	34
6	Výsledky třesu	34
7	Výsledky naklonění	34

1 Úvod

Automobilky se po několik desetiletí vždy snažily a snaží zdokonalovat automobily všelijakými způsoby, ať už se jedná o materiál karoserie, pohodlí posádky, její bezpečnost, výkon aut a v mnoha dalších směrech. Tato bakalářská práce se zabývá jedním z těchto směrů, a to bezpečností řidičů v automobilech. Detekce řidiče může mít zásadní význam pro bezpečnost posádky, neboť se může automaticky zaznamenat viditelný špatný stav řidiče a podle toho tak jednat. Pro lidské oko lze celkem jednoduše zjistit podle náznaků v oblasti hlavy, zda se dané osobě něco neděje popřípadě nestalo, avšak pro strojovou techniku je to mnohem složitější. Strojová technika vyžaduje úpravy obrazu do forem, s kterými různé algoritmy dokáží pracovat a následně vyhodnocovat to, co člověk jako takový dokáže zhodnotit skoro instantně.

Práce je zaměřená na rozpoznávání nepříznivých příznaků člověka pomocí zájmových bodů obličeje. Pro získání těchto zájmových bodů slouží obyčejná kamera s využitím již známých algoritmů.

V teoretické části jsou rozebrány technologie, které jsou v dnešní době používány a vysvětleno jak pomáhají s rozeznáváním lidských tváří. Také je popsán celkový postup práce, výhody a nevýhody využitých metod.

V praktické části je představen návrh a realizace programu pro detekci a následné analýzy lidských tváří. Navržené algoritmy jsou testovány na videích s rozdílným úhlem pohledu a následně porovnávány. V poslední řadě je zrekapitulovaný celkový směr práce a jsou nastíněny další možnosti jak práci rozšířit.



Obrázek 1: Frontální snímání tváře řidiče. [1]

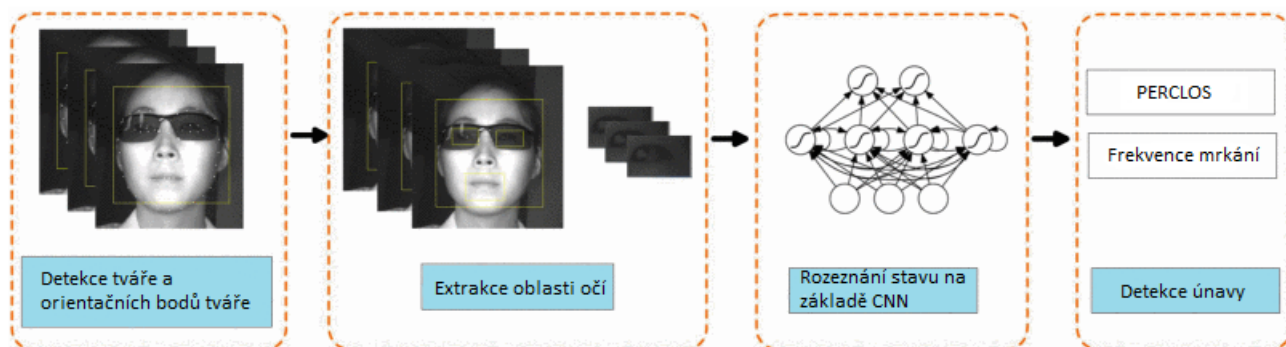
2 Detekce tváře

Lidské tváře obsahují mnoho informací, podle kterých se dá analyzovat stav zdravotního stavu. Sledováním důležitých částí tváří můžeme odhalit například únavu řidiče, díky které jsou nejčastější dopravní nehody, nebo i odhalení asymetrie výrazu člověka, která může signalizovat problém v mozku. Musíme si však uvědomit, že žádná osoba nemá stejné obličejové rysy. Okolní podmínky, zejména světlo, prach atd. tomu nepomáhají a části obličeje můžou být překryté třeba slunečními brýlemi a v neposlední řadě určitou obtíž je i pozorovací úhel. Za několik posledních desetiletí byly vytvořeny různé metody pro detekci tváří, každá metoda má však své plusy a mínusy.

I přes to, že tato oblast je zkoumána po více než několik desetiletí, stále neexistují metody, které by fungovaly se stoprocentní úspěšností.

V 60. letech minulého století vědci W. Bledsoe, H. Chan, Wolf a Ch. Bisson vyvinuli první systém pro rozpoznávání tváře. Jejich program vyžadoval aby správce lokalizoval základní rysy a poté se vypočítaly vzdálenosti a poměry ke společnému referenčnímu bodu, který se pak porovnal s referenčními údaji.[2]

P. Weyers, A. Barth a A. Kumment ve své práci, kde se zaměřují pouze na únavu řidiče, používají pro detekci algoritmus AdaBoost se souborem lokálních binárních znaků (LBF) pro získání zájmových bodů obličeje. Poté se rozeznává únavu na základě konvulční neuronální sítě (CNN), postup je zobrazený na obrázku 2. I díky tomu, že používají infračervené videa, jejich výsledky ukazují, že jejich metoda poskytuje velmi přesné a velmi robustní rozpoznávání. Více v článku: [3]



Obrázek 2: Blokové schéma. [3]

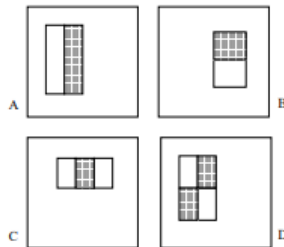
V současné době jedna z nejvíce známých metod, která může být použita k detekci obličeje, je metoda s využitím Haarových příznaků. V článku[4] je metoda stručně vysvětlena s ukázkou implementace a výsledkem.

2.1 Haarovy příznaky

Detekce objektů založených na Haarových příznacích je efektivní metoda detekce objektů navržená Paulem Violou a Michaellem Jonesem [5]. Jedná se o přístup založený na strojovém učení, kde kaskádová funkce je vycvičena z mnoha pozitivních a negativních snímků. Pro dosažení vysokých rychlostí tento systém vyžaduje převod rastrového obrazu do šedé stupnice.

2.1.1 Detekce příznaků

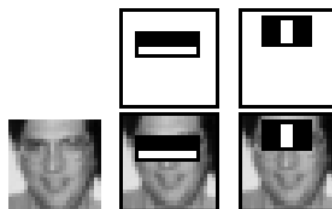
Základní princip Haarových příznaků spočívá v rozdílu součtu pixelů, co se týče stupně šedi ve dvou sousedících obdelníkových oblastech. Čím větší výsledný rozdíl bloků je, tím větší šance, že se v této oblasti nachází haarův příznak. Tyto oblasti jsou stejně velké, mají stejný tvar a můžou být vertikálně nebo horizontálně oddělené, viz obrázek 3. Pro zrychlení výpočtů příznaků se používá převod oblastí do integrální podoby, to znamená že každé pole v oblasti je reprezentováno jako součet hodnot všech pixelů směrem vlevo a nahoru.



Obrázek 3: Typy příznaků (převzato z [5])

2.1.2 Klasifikace nalezených příznaků

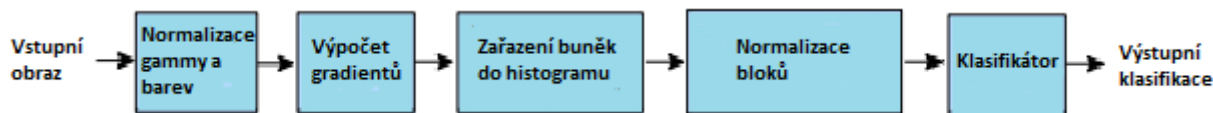
Ke klasifikaci všech nalezených příznaků se používá algoritmus Adaptive Boosting, zkráceně AdaBoost. Tento algoritmus sestavuje finální klasifikátor z množiny nalezených silných příznaků (ukázka nalezených příznaků na obrázku 4). AdaBoost iterativně prochází trénovací obrazy, při každé iteraci se zvyšuje váha chybně klasifikovaných obrazů. Průběh se opakuje, dokud není dosaženo přijatelné přesnosti. Konečný klasifikátor je tedy součet slabých klasifikátorů. Slabé klasifikátory samotné nedokážou klasifikovat pořádně celý objekt, ale se součtem těchto klasifikátorů přichází vysoká úspěšnost.



Obrázek 4: První dva příznaky vybrány AdaBoostem Převzato z: [5]

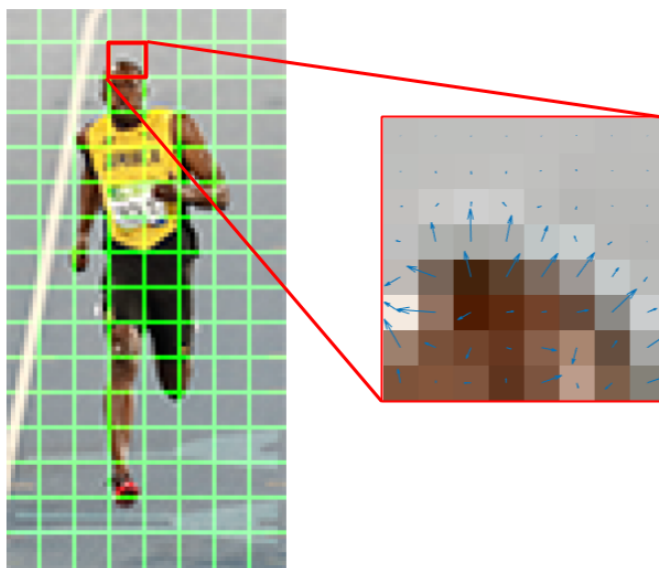
2.2 Histogram orientovaných gradientů (Histogram oriented gradients)

S další metodou přicházejí autoři Sh. Chang, D. Xiaoqing a F. Chi založenou na histogramu orientovaných gradientů (HOG). HOG byl sice původně vytvořený za účelem detekce lidských postav, nicméně autoři v článku [7] se zaměřili na detekci tváře. Jejich výsledky ukazují kvalitní míru detekce s nízkým výpočtovým časem.



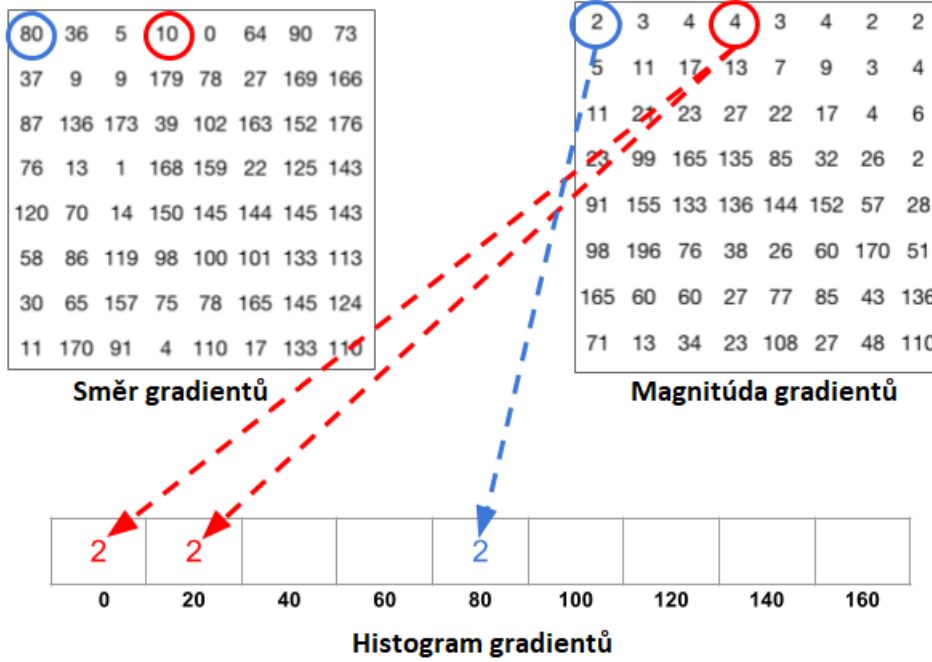
Obrázek 5: Blokové schéma HOG (převzato z [6])

Histogram orientovaných gradientů (HOG) je deskriptor vlastností, který je široce používán k charakterizaci objektů prostřednictvím jejich tvarů. Vzhled a tvar hledaného objektu lze často popsat rozložením lokálních gradientů intenzity nebo směrů hran. HOG je široce využíván jako detektor objektů, jako jsou například vozidla, dopravní značky, lidské tváře a mnoho dalších objektů.



Obrázek 6: Zobrazení gradientů každého pixelu jedné buňky (převzato z [8])

Princip této metody spočívá ve vypočítávání gradientů, tedy velikost a směr pro každý pixel (ukázka velikosti a směru gradientů na obrázku 6). Pixely se následně seskupují do buněk určité velikosti, například 8x8 pixelů pro každou buňku. Pro zvýšení efektivity vyhledávání objektů by měly být ještě předem normalizovány barvy obrazu a svítivost, protože různé obrazy mohou mít různý kontrast. Na obrázku 7 lze vidět, že histogram se rozděluje do devíti různých kanálů, toto je však pouze příklad, neboť počet kanálů může být libovolný. Podle orientace gradientu každého pixelu se zařadí do daného kanálu jeho velikost, výsledkem kanálu je součet těchto velikostí.



Obrázek 7: Přiřazování gradientů do histogramu (převzato z [8])

V poslední řadě se buňky zformují do jednotlivých bloků, například 1 blok = 2x2 buňky. Tyto bloky se částečně překrývají, takže buňky mohou být součástí více bloků. Jelikož každá buňka má právě jeden histogram, součástí jednoho bloku jsou tedy 4 histogramy. Tyto histogramy jsou zřetězené do jednorozměrného vektoru o 4x9 hodnotách. Vektory jednotlivých bloků jsou zřetězené a tato podoba je předána do finálního klasifikátoru, například SVM který je popsáný v jednom z následujících částí práce. Tento postup je znázorněný v blokovém schématu obrázku 5.

2.2.1 Výpočet gradientů

Při zpracování obrazu potřebujeme vědět, jakým směrem se mění barvy. Proto chceme měřit gradient na pixelech barev. Vektor gradientu je definován jako metrika pro každý jednotlivý pixel, který obsahuje změny barev pixelů v ose x i ose y. Vektor gradientu je daný jako:

$$\nabla f(x, y) = \begin{bmatrix} g_x \\ g_y \end{bmatrix} = \begin{bmatrix} \frac{\partial f}{\partial x} \\ \frac{\partial f}{\partial y} \end{bmatrix} = \begin{bmatrix} f(x+1, y) - f(x-1, y) \\ f(x, y+1) - f(x, y-1) \end{bmatrix}$$

Termín $\delta_f \delta_x$ je parciální derivace ve směru x, který je počítán jako rozdíl barev mezi sousedními pixely na levé a pravé straně cíle, $f(x+1, y) - f(x-1, y)$. Podobně, termín $\delta_f \delta_y$ je parciální derivace na y-směr, tedy: $f(x, y+1) - f(x, y-1)$. Gradient má 2 základní atributy: Směr a magnitudu.

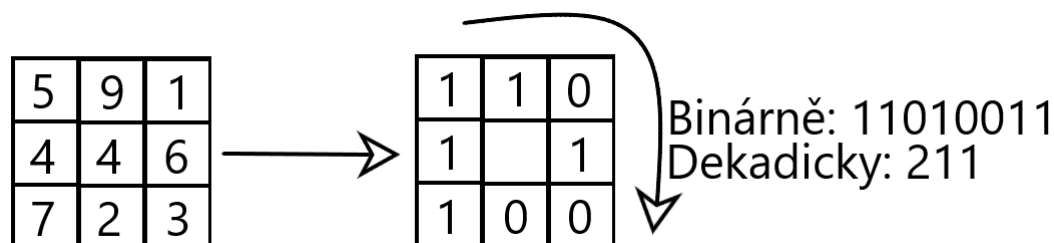
Vzorec pro výpočet směru: $\Theta = \arctan(g_y/g_x)$

Vzorec pro výpočet magnitudy: $g = \sqrt{g_x^2 + g_y^2}$

2.3 Local Binary Pattern

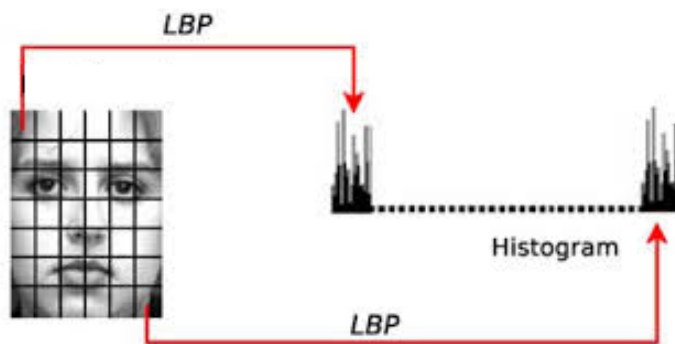
Další metoda pro detekci tváře, která může být použita, je metoda založená na principu lokálního binárního vzoru. Metoda byla původně zveřejněna za účelem detekce textur, nicméně po letech našla uplatnění i v detekci tváře, jako to například ukázali autoři T.Ahonen, A.Hadid a M.Peitikainen v článku [9].

Principem je rozebírání pixelů do 3x3 bloků, přičemž je snaha o reprezentace bloku jedinou hodnotou. Všechny 8 okrajových pixelů je porovnáváno se středním pixelem. Porovnávané hodnoty jsou ve formě stupnic šedi. Pokud je hodnota krajního pixelu menší, než středního pixelu, na pozici tohoto krajního pixelu se nastaví binární hodnota 1, v opačném případě 0. Tyto hodnoty pak reprezentují jedno bajtové číslo, které se pak převede do dekadické formy, viz obrázek 8. Velkou výhodou je, že výsledek se nemění se změnou světla.



Obrázek 8: Převod blok pixelů do dekadické hodnoty (převzato z [9])

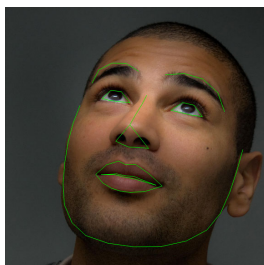
Tímto způsobem z každého bloku získáme číslo v rozsahu 0 až 255. Tyto hodnoty každého bloku se uloží do histogramu, jako je to na obrázku 9. Počítáním množství stejných výsledků se dá předpovědět, co na obrázku vidíme.



Obrázek 9: Ukládání počet stejných výsledků do gradientu (převzato z [9])

3 Zájmové body tváře

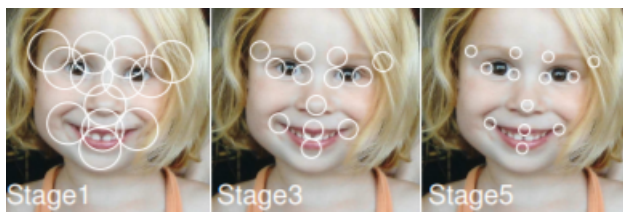
Důležitou součástí této práce je detekce zdravotního stavu řidiče, to je dosaženo pomocí získávání zájmových bodů obličeje. K lokalizaci těchto bodů je nejprve potřeba detekce tváře, jako to bylo popsáno v předchozí kapitole. Pokud je detekce úspěšná, může se přejít na další krok, a to k nalezení těchto bodů. Do dnešní doby bylo prezentováno mnoho metod, kde každá má jiný princip řešení problému, jinou rychlost vypočítávání a počet nalezených orientačních bodů se může lišit. Algoritmy detekce orientačních bodů obličeje se rozdělují do tří hlavních kategorií. Yue Wu a Qiang Ji [10] ve své práci píší o základních principech těchto kategorií, porovnávají rozdíly, výhody a nevýhody. Do první kategorie spadají holistické metody, které explicitně vytvářejí modely, které představují globální informace o vzhledu a tvaru. Další metodou je takzvaný omezený lokální model, který vytváří pohled lokálního modelu. Do poslední kategorie spadá metoda založená na regresi. Ta implicitně zachycuje informace tvaru a vzhledu obličeje. Ukázka ohraničení tváře na základě zájmových bodů je na obrázku 10.



Obrázek 10: Dřív zájmové body (převzato z [11])

3.1 Regrese lokálních binárních znaků

V práci Face Alignment at 3000 FPS via Regression Local Binary Features [12] autoři Sh. Ren, Xudong Cao, Yichen Wei, Jian Sun používají regresní přístup pro obličejové rysy s využitím lokálních binárních znaků a princip lokality. Metoda dosahuje velice dobré rychlosti. Princip lokality vede k naučení souboru vysoce diskriminovaných lokálních prvků nezávislé na sobě. Získané lokální binární prvky se používají k učení lineární regresi pro finální výstup. Obličejová forma je odhadovaná a vylepšovaná kaskádově, viz obrázek 11.

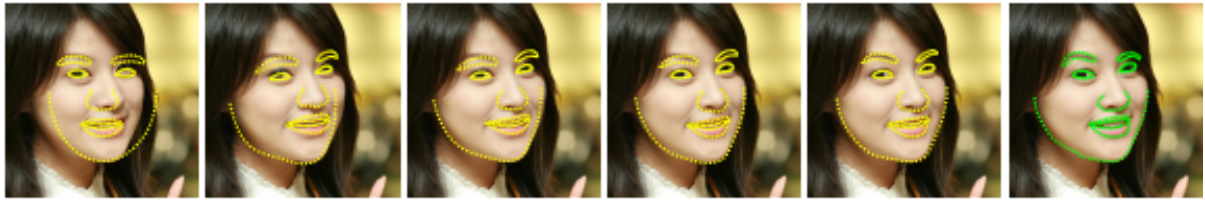


Obrázek 11: Vylepšování bodů kaskádově (převzato z [12])

3.2 One Millisecond Face Alignment

Tento algoritmus dosahuje výsledků v milisekundách a jeho přesnost je vysoká. Tento algoritmus představili V.Kazemi a J.Sullivan[13]. Nevýhodou je, že potřebujeme spolehlivé funkce pro přesné předpovídání tvaru a na druhé straně potřebujeme přesný odhad tvaru pro získání spolehlivých funkcí. K řešení tohoto problému se využívá kaskádového přístupu.

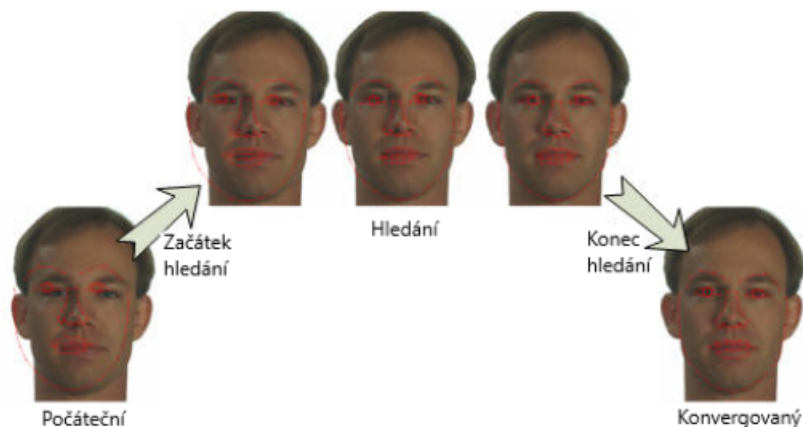
Dále namísto odhadování parametrů tvaru na základě vlastností extrahovaných z globálního souřadného systému obrazu, je obraz nejprve převeden do normalizovaného souřadnicového systému a následně odhadované parametry tvaru jsou získávány z tohoto souřadnicového systému. Tento proces se obvykle několikrát opakuje až do konvergence, jako to je na obrázku 12.



Obrázek 12: Zlepšování odhadu na základě regresních stromů (převzato z [13])

3.3 Aktivní model tvarů (ASM)

T.F.Cootes v roce 1995[14] přišel se metodou aktivních modelů tvarů. U této metody se z určité množiny, ve které jsou manuálně označené významné body obličeje (oči, rty, atd.) vygeneruje normalizovaný model vzhledu a tvaru. Při zpracování nového obrazu se ASM snaží pomocí iterčního přístupu rozmístit body normalizovaného modelu tak, aby nejlépe odpovídal skutečným místům těchto bodů v analyzované tváři (viz obrázek 13). Při stanovení konvergence získáme konečný výsledek.



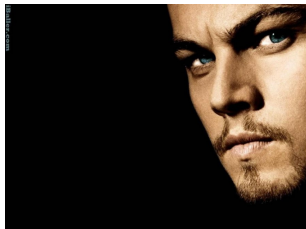
Obrázek 13: Ukázka ASM (převzato z [15])

3.4 Datové sety

Existuje několik datasetů, které obsahují trénovací data speciálně vytvořené pro detekci zájmových bodů tváře. Tyto datasety jsou využívány pro trénování jednotlivých detektorů. Každý datový set může mít jiný počet obrázků, jiné manuální označení pozic a jiné testovací obrazy. V této sekci jsou některé datasety popsány.

3.4.1 300-W

Ibug 300-w[16] je datový set, který byl vytvořený za účelem soutěže. Jedná se o datový set, který obsahuje 300 venkovních snímků a 300 vnitřních snímků. Trénovací set je kombinací automatické metodiky a ručním zásahem anotován 68 orientačními body. Ukázky obrázků datové sady 14(a), 14(b).



(a)



(b)

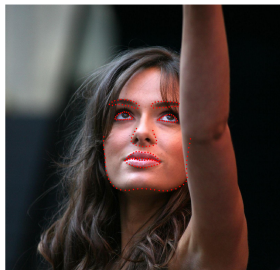
Obrázek 14: Ukázka obrázku z 300-W sady [16]

3.4.2 HELEN

Datový set HELEN [17] byl sestaven z obrázků na Flickr, při každém hledání bylo využito slovo "portrait" s rozšířením jako třeba "outdoor". Autoři se snažili, aby datový set byl složen z nejrozličnějších příkladů s podmínkou vysokého rozlišení. Obrazy byly ručně označovány (194 bodů) a výsledný set obsahuje 2000 trénovacích a 330 testovacích obrázků s vysoce přesnými, detailními a konzistentními anotacemi primárních komponent obličeje, viz obrázky 15(a), 15(b).



(a)

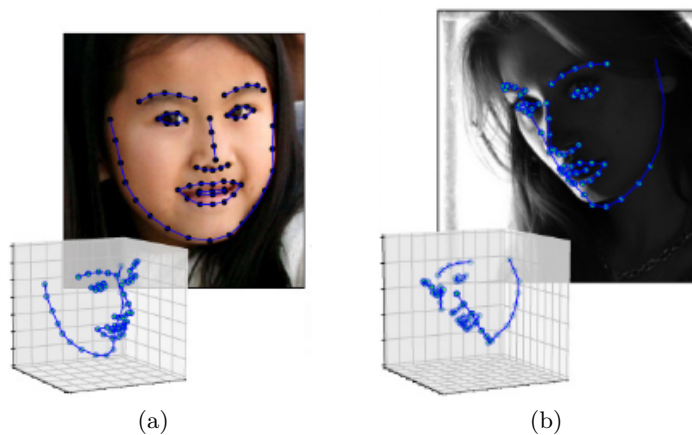


(b)

Obrázek 15: Ukázka obrázků z HELEN sady [17]

3.4.3 3DFAW

Nejběžnějším přístupem je zarovnání, které považuje tvář za 2D objekt. Tento předpoklad platí tak dlouho, dokud je obličej čelní. Vzhledem k tomu, že se orientace obličeje může lišit od frontální orientace, tento předpoklad se rozpadá. Datový set "3D Face Alignement in Wild"[18] vznikl za účelem spojit výzkumníky, jejichž práce souvisí s 2D a 3D zarovnáním obličeje. Datový soubor 3DFAW obsahuje přes deset tisíc skutečných a syntetických obrazů obličeje s 66 3D anotacemi. Ukázka obličeje s 3D anotacemi na obrázcích 16(a), 16(b).



Obrázek 16: Ukázka obrázku z 3DFAW sady [18]

3.4.4 LFPW

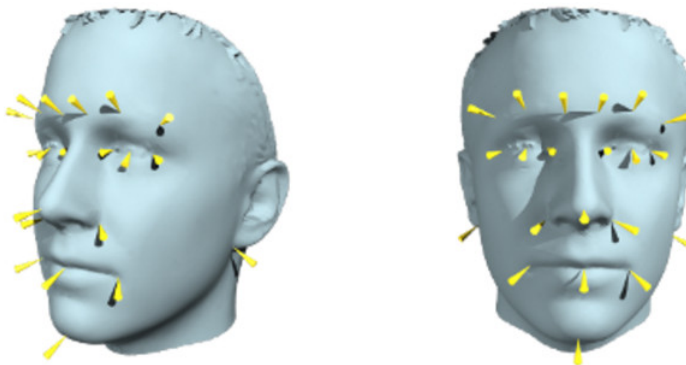
LFPW neboli "Labeled Face Parts in the Wild"[19] je datový set vytvořený společností Kriegman-Belhumeur Vision Technologies. Tento set se skládá z 1433 tváří stažených z google.com, flickr.com a yahoo.com. Každý obraz byl označen 29 anotacemi, viz obrázek 17. Set se rozděluje na 1133 trénovacích a 300 testovacích obrazů.



Obrázek 17: Ukázka anotace LFPW. Převzato z: [19]

3.4.5 AFLW

Datový set "Annotated Facial Landmarks in the Wild"[20] poskytuje rozsáhlou sbírku barevných obrazů obličejů shromážděných z Flickr, které vykazují rozsáhlý rozměr výrazu, póz, věku, pohlaví (41% mužů, 59 % žen). Celkově je anotováno 25 000 tváří až s 21 body na obraz. Celkově datový set obsahuje kolem 380 tisíc manuálně označených bodů, k dispozici není žádná anotace pokud není viditelný orientační bod, například při pootočení hlavy není anotován jeden z ušních laloků, viz obrázek 18. Tento set byl vytvořený univerzitou Graz University of Technology.



Obrázek 18: Ukázka anotace AFLW. Převzato z: [20]

3.4.6 COFW

Datový set COFW [21] je vytvořený institutem California Institute of Technology. Tento dataset se liší od ostatních zejména tím, že tváře v obrazech jsou velmi často překryté brýlemi, rukou, čepicí atd. Datový set obsahuje 1345 barevných i černobílých obrazů pro trénování a 507 obrazů pro testování a 29 anotacemi tváře. Ukázka obrázků této sady v obrázku 19.

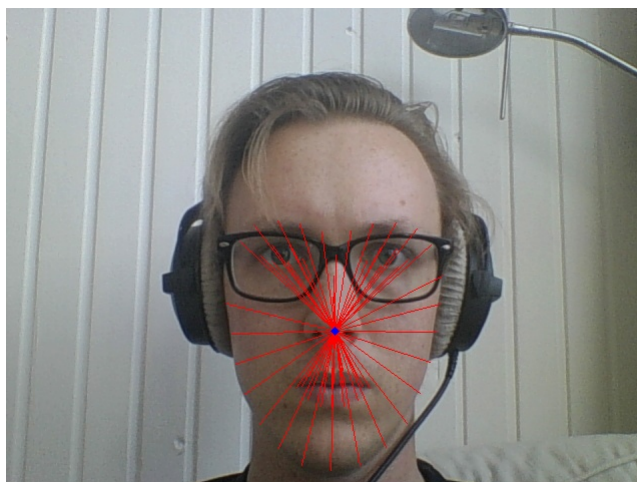


Obrázek 19: Ukázka COFW. Převzato z: [21]

4 Klasifikace akcí

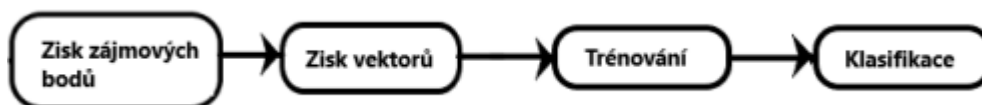
Důvody, proč byly všechny předchozí části popsány, směřují k tomuto segmentu. Nyní když jsou k dispozici zájmové body tváře, lze díky nim určit aktuální stav řidiče. Existuje mnoho algoritmů, které zpracovávají zájmové body různým způsobem. Jednou z možností je měřit vzdálenosti nebo úhly těchto bodů. Například kdyby bylo úkolem zjistit, zda se člověk usmívá nebo ne, teoreticky stačí měřit vzdálenost krajních bodů pusy. Samozřejmě existují metody, které můžou být více komplexní, které však vykazují velice dobré výsledky.

Například autor článku [22] se ve své práci snažil rozeznávat emoce lidí. Ze všech získaných zájmových bodů vypočítal těžiště, od které pak měřil vzdálenost, viz. obrázek 20. Tyto vzdálenosti použil jako trénovací data pro lineární klasifikátor SVM.



Obrázek 20: Vzdálenosti bodů od těžiště [22]

Na druhou stranu Fuzail Khan [23] zase ve svém dílu použil zájmové body tváře pro natrénování neuronové sítě. Měřením vzdáleností a směru bodů získal vektorový formát, který použil jako trénovací množinu (viz obrázek 21). Výsledky v jeho práci ukazují velice dobrou úspěšnost.



Obrázek 21: Blokové schéma klasifikace. Inspirováno z: [23]

4.1 SVM

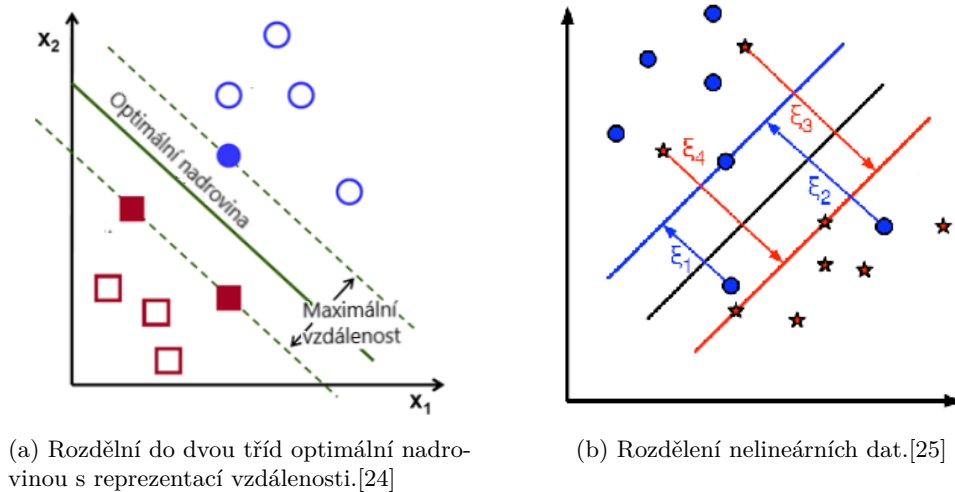
Jak bylo řečeno na předchozí straně, metoda SVM může být použita pro práci se zájmovými body. SVM neboli metoda podpurných vektorů je model řízeného učení, určený především pro klasifikaci a regresi. Support vector machine umožňuje najít takovou nadrovinu, která dokáže rozdělit lineárně i nelineárně separovatelnou úlohu na dvě třídy. Dále SVM využívá jádrové transformace, která dokáže převést nelineárně separovatelná data na lineárně separovatelná. [24]

4.1.1 Lineární SVM

Jak už samotný název napovídá, lineární SVM je určený pro data, které jsou lineárně separovatelná. Nalezení optimální nadroviny může být klíčové pro správné fungování rozeznávání dat. Ve 2D prostoru je optimální nadrovinou přímka, ve 3D prostoru to je zase rovina. Optimální separaci považujeme nalezení nadroviny, kde vzdálenosti bodů od nadroviny jsou co nejvyšší, jako je tomu na obrázku 22(a).

4.1.2 Nelineární SVM

Většina vzorků, s kterými se pracuje nelze jednoduše rozdělit do lineárně separovatelných dat a vyžaduje výkonnější nástroj než jednoduchý lineární klasifikátor. Díky tomu, že data jsou nelineárně oddělitelná, některé případy se budou chybně klasifikovat. To znamená, že je potřeba najít takovou nadrovinu, která nebude umožňovat velkou chybovost, a zároveň vzdálenost podpurných vektorů od této nadroviny musí být pořád co největší, viz obrázek 22(b).



Obrázek 22: Ukázka lineárně a nelineárně rozdělitelných dat

5 Vlastní implementace a testování

Jak už bylo řečeno, úkolem práce je detekce zdravotního stavu řidiče. S využitím kamery a zájmových bodů je možno zanalyzovat vstupní obraz a zjistit důležité informace o změnách v obličeji řidiče(například zavřená nebo otevřená ústa) a kombinací těchto informací lze vygenerovat na výstup programu aktuální stav řidiče.

Například únava se projevuje mnohými způsoby. Průměrná frekvence mrkání je 5 až 15 mrkutí za minutu. Člověk, který je unavený mrká méně než člověk, který je čilý. Naopak časté mrkání může mít příčinu neurologickou nebo i psychologickou. Dalším příznakem únavy je časté zívání. Také je potřeba zmínit, že s únavou může nastat mikrospánek. O mikrospánku můžeme spekulovat, pokud osoba má zavřené oči alespoň určitou dobu, tedy 3 až 15 vteřin, poté dojde k probuzení nebo usnutí. V této práci jsem se zaměřil na několik těchto stavů a to:

- Mrknutí
- Zívnutí
- Mikrospánek
- Třes hlavy
- Nakloněnní hlavy
- Nesymetrie tváře

K otestování jsem vytvořil aplikaci ve vývojovém prostředí Microsoft Visual Studio 2017 v jazyce C++ s knihovnami OpenCV a Dlib. Knihovna Dlib nabízí kompletní detektor předních tváří s využitím deskriptoru HOG, který jsem využil. Pro získání zájmových bodů tváře jsem použil prediktor tvaru využívající datový soubor shape-predictor-68-face-landmarks.dat trénovaný na datové sadě w-300.

5.1 Použité knihovny

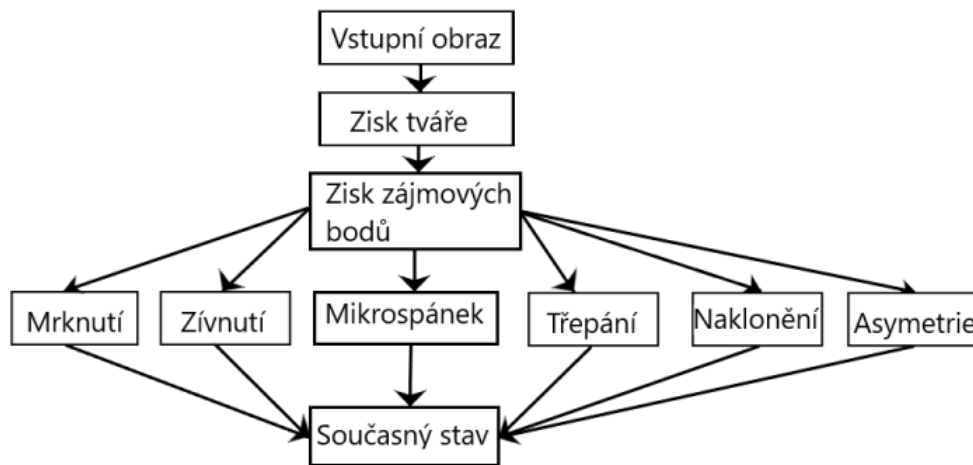
OpenCV [26] je volně dostupná knihovna, která se zaměřuje na odvětví výpočetní matematiky strojového vidění i na zpracování obrazu v reálném čase s umožněným využitím pro více jádrové zpracování. Knihovna zahrnuje více než 2500 optimalizovaných algoritmů pro práci se zpracováním obrazu a strojového vidění a byla původně vyvíjena společností Intel. OpenCV bylo využito u mnohých projektů, například při navigaci robotů, monitorování bazénů kvůli prevenci tonutí lidí, kontrolování štítků u produktů a mnoho jiných projektů. Tuto knihovnu využívají známé společnosti jako jsou Google, Yahoo, Microsoft, Intel, IBM a mnoho dalších firem. Knihovna podporuje platformy Windows, MacOS, Android a Linux a jazyky C++, Python, Java a MATLAB.

Dlib [27] je moderní volně dostupná knihovna s algoritmy a nástroji pro strojové učení v C++. Vývoj probíhá od roku 2002 a již v roce 2016 obsahoval softwarové komponenty pro práci

s vlákny, datových struktur, lineární algebrou, strojového učení a zpracování obrazu a další. Používá se jak v průmyslu, tak i v akademické sféře a to v oblasti robotiky, telefonů nebo i v prostředí s velkým výpočetním výkonem.

5.2 Postup práce

Postup práce, který jsem zvolil se dělí do tří základních částí. V první části implementace jsem se zabýval zpracováním obrazu z kamery a následné extrakce tváře, další část je zaměřená na získání zájmových bodů tváře. Poslední část se zabývá vypočítáváním na základě informací získaných zájmových bodů a vygenerovávání výsledku aktuálního zdravotního stavu řidiče, blokové schéma práce je zobrazené na obrázku 23.



Obrázek 23: Blokové schéma implementace.

5.2.1 Zpracování obrazu a získání tváře

Společným krokem pro celou analýzu rozpoznání stavu řidiče je detekce tváře. K detekci tváře je použita knihovna OpenCV s kombinací knihovny Dlib. Nejprve ale bylo zapotřebí získat obraz z kamery, k tomu slouží třída `cv::VideoCapture` s metodou `open()`. Do této metody se posílá parametr jako index kamery. Následně do metody `read()` se zašle parametr `temp`, do které se uloží data z obrazu v maticové formě. Ukázka je ve výpisu 1.

```
cv::VideoCapture cap;  
cap.open(0);  
cv::Mat temp;  
cap.read(temp);
```

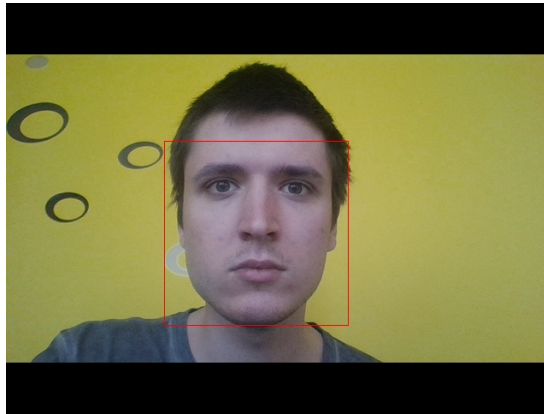
Výpis 1: Získání obrazu z kamery

V této fázi mám obraz z kamery uložený v proměnné `temp`. Teď už stačí na tomto obrazu detekovat tvář. K tomu přispívá metoda z knihovny Dlib `get_frontal_face_detector()`, která

vrací detektor tváří. Ještě než jsem mohl poslat obraz do detektoru, který vrací nalezené tváře do vektorového pole, bylo nutné převést matici obrazu z reprezentace OpenCV do DLib (postup ve výpisu 2). Příklad výsledku těchto kroků je na obrázku 24, kde získaná tvář je ohraničená v červeném rámcí.

```
frontal_face_detector detector = get_frontal_face_detector();  
cv_image<bgr_pixel> cimg(temp);  
std::vector<dlib::rectangle> faces = detector(cimg);
```

Výpis 2: Získání tváří z obrazu



Obrázek 24: Získaná tvář ohraničená v červeném rámcí.

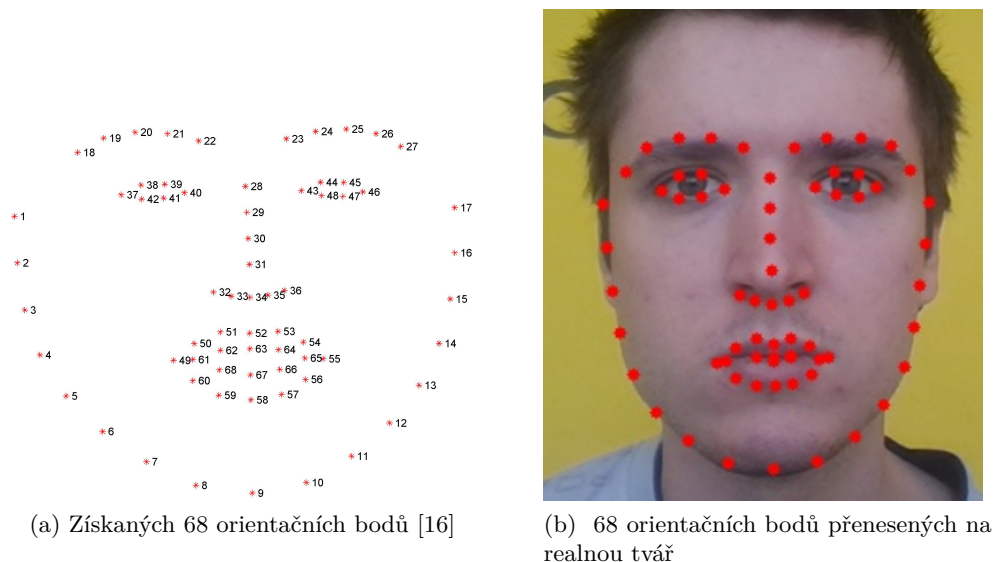
5.2.2 Získání zájmových bodů

V případě, že se v předchozím kroku úspěšně našla nějaká tvář, je nutné pro mou práci z ní získat zájmové body. Ty získám pomocí predikátoru tvaru objektu z knihovny Dlib, do které načítám datový model ze souboru `shape_predictor_68_face_landmarks.dat`, viz výpis 3. Model je natrénovaný na datové sadě 300-W, tento datový soubor je volně dostupný na oficiálních stránkách Dlib [28].

```
shape_predictor pose_model;  
// nahrání modelu  
deserialize("shape_predictor_68_face_landmarks.dat") >> pose_model;  
std::vector<full_object_detection> shapes;  
for (unsigned long i = 0; i < faces.size(); ++i) {  
    shapes.push_back(pose_model(cimg, faces[i]));  
}
```

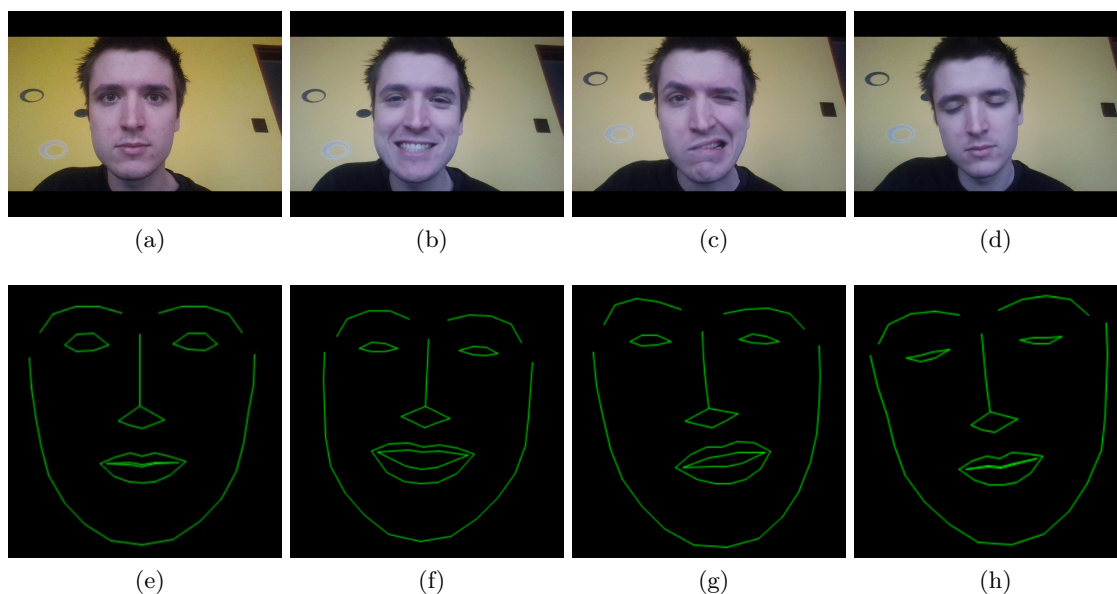
Výpis 3: Získání orientančních bodů

Každá výsledek predikce tvaru tváře je uložena ve vektorovém poli shapes. Každý element vektorového pole shapes obsahuje 68 bodů, kde každý bod reprezentuje souřadnici ve 2D zobrazovacím systému. Tyto body jsou indexované podle obrázku 25(a), například pro levé obočí platí rozptyl od bodu 18 až po 22. Ukázka zájmových bodů na reálné tváři v obrázku 25(b).



Obrázek 25: Reprezentace 68 zájmových bodů.

Ukázky ohraničení tváře na základě zájmových bodů v obrázcích 26(a), 26(b), 26(c), 26(d), 26(e), 26(f), 26(g), 26(h).



Obrázek 26: Různé pohledy a viditelné změny bodů

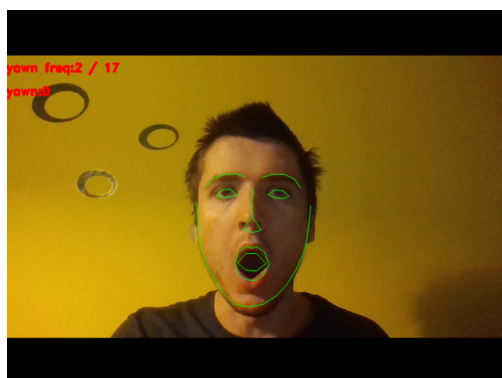
V případě, že byla nějaká tvář nalezena a byly získané zájmové body této tváře, mohou vypočítávat aktuální stav řidiče pomocí souřadnic bodů. V následující části budou rozebrané implementované metody, kterými jsem určil aktuální stavy, které byly zmíněny v popisu problému.

5.3 frekvence mrkání, zívání a mikrospánek

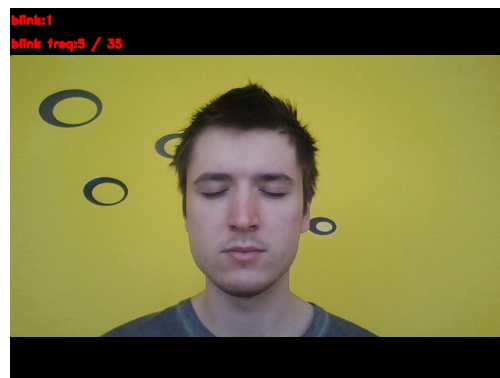
První metoda kterou si rozebereme je zaměřená na mrkání. K problému detekce frekvence mrkání jsem přistoupil s následujícím přístupem. Na základě experimentu jsem vypočetl, že jedno z vhodných řešení je měřit vzdálenost vrchních víček od spodních (body 39 a 41, body 44 a 48). Tuto hodnotu jsem následně porovnával s hodnotou $1/19$ vzdálenosti zájmových bodů 28 a 9. Pokud výsledná hodnota byla nižší, algoritmus vyhodnotil akci jako mrknutí.

Stejný princip jsem uplatnil při testování zívání. Na základě zkoušení detekce zívání jsem měřil vzdálenost spodního od horního rtu, tedy body 52 a 58. K porovnání hodnot jsem opět použil vzdálenost zájmových bodů 28 a 9. Nejpřesnějších výsledků bylo dosaženo porovnáváním vzdálenosti mezi spodním a horním rtem s hodnotou $1/3.3$ vzdálenosti bodů 28 a 9. Pokud výsledná hodnota byla vyšší, jednalo se o zívnutí.

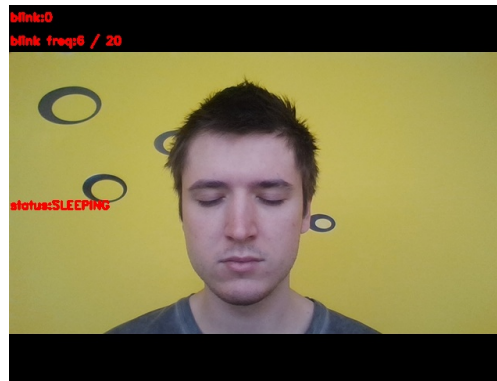
K dosažení stavu mikrospánku jsem z předchozích poznatků implementoval metodu, která po 3 vteřinách zavřených očí říká, že daná osoba spí. Problémem může být, kdy se chybně klasifikuje uprostřed tohoto intervalu mrknutí, takže zabere další 3 vteřiny, než se potvrdí, že daná osoba spí. Ukázky mrknutí, zívnutí a mikrospánek jsou na obrázcích 26(a), 26(b) a 26(c).



(a) Zívnutí



(b) Zaznamenání mrknutí



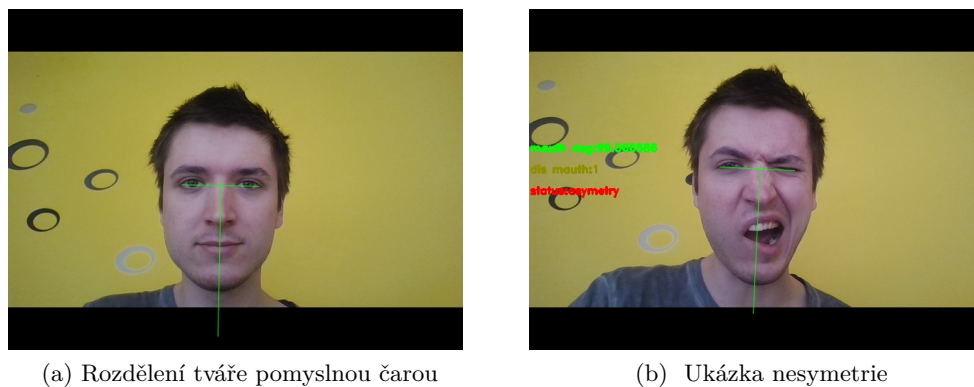
(c) Mikrospánek [16]

Obrázek 26: Ukázka zívnutí, mrknutí, mikrospánek

5.4 klasifikace asymetrie

Nesymetrie ve tváři může signalizovat mnoho příznaků, od jakékoli bolesti až po poruchu svalové funkce. Na základě zkoumání jsem vytvořil dvě metody, které jsem později porovnal.

U první metody jsem tvář rozdělil na dvě poloviny pomyslnou čarou a od ní jsem měřil vzdálenosti bodů. K tomu jsem využil krajní body očí, jelikož tyto body jsou jedny z mála, které jsou statické. Přes tyto dva body jsem určil přímku a od ní jsem pomocí normálového vektoru od středu kořene nosu vytvořil přímku, které dělí tvář na dvě symetrické poloviny. Následně jsem měřil úhly a vzdálenosti koutků úst od půlící čáry, tyto hodnoty jsem pak porovnával, zda se hodnoty liší. Ke klasifikaci nesymetrie jsem také využil obočí, kde jsem získal vzdálenosti obočí od očí, které jsem pak porovnával. Ukázka rozdělení tváře je na obrázku 27(a) a tvář, která je vyhodnocená, jako nesymetrická je na obrázku 27(b).



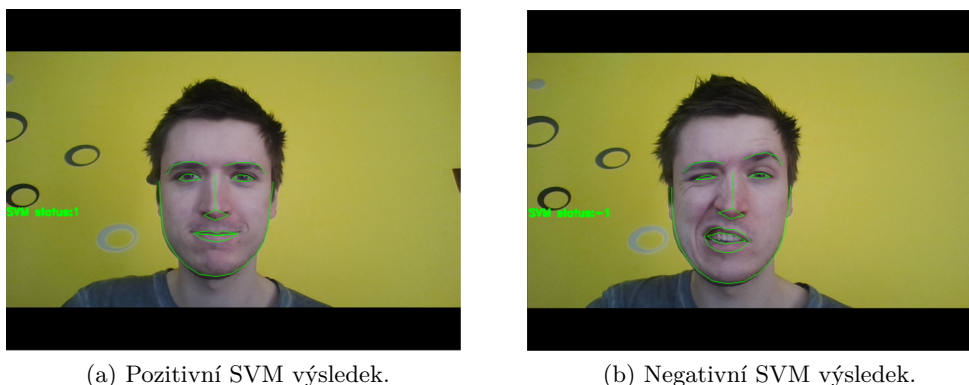
(a) Rozdělení tváře pomyslnou čarou

(b) Ukázka nesymetrie

Obrázek 27: Asymetrie první metody

Druhý postup je založený na lineárním klasifikátoru SVM, který podporuje knihovna OpenCV. Jelikož jsem nenašel žádný datový set pro nesymetrie tváře člověka, musel jsem si takový set vytvořit. V tomto případě jsem tyto data vytvořil na sobě. Data jsem rozdělil do dvou skupin,

kde obě z této skupiny mají 500 trénovacích vzorků. První skupina jsou data, kdy se člověk tváří normálně, mluví, usmívá se. Ve druhé skupině se nachází data, kdy pohled člověka je abnormální, nesymetrický. Pro každý obraz jsem získal zájmové body tváře, mezi několika těchto bodů jsem změřil a uložil vzdálenosti do datového souboru (vzdálenosti bodů mezi 25 a 52, 20 a 52, 42 a 49, 48 a 55, 39 a 41, 44 a 48, 50 a 57, 54 a 59), se kterým se později tento klasifikátor naučil. Ukázka výsledků této metody na obrázcích 28(a) a 28(b).



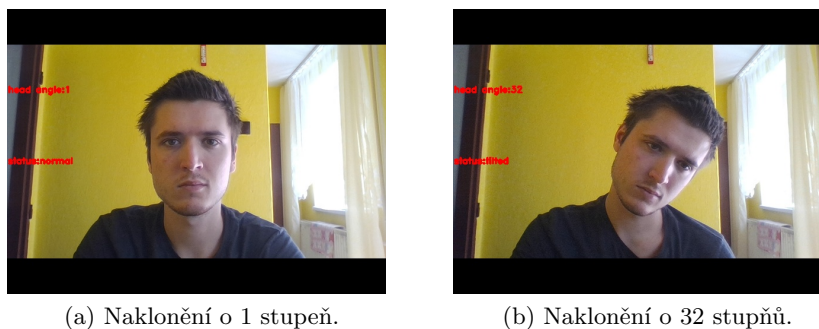
Obrázek 28: Asymetrie druhé metody

5.5 Třes

Třes hlavy je další nechtěnou situací. Taková situace může například signalizovat epileptický záchvat. Na základě pokusů jsem vytvořil algoritmus, který pracuje tak, že měří vzdálenosti mezi dvěmi zájmovými body na obvodu tváře (1 a 17) od špičky nosu. Tyto hodnoty se vzájemně porovnávají na po sobě jdoucích snímcích. Pokud jsou získané hodnoty mezi posledními a aktuálními snímky vysoké, poukazuje to na nestandardní rizikový stav.

5.6 Naklonění

Nakloněná hlava o velký úhel může zase znamenat, že osoba upadla do bezvědomí. Pro správné vyhodnocení je důležité umístění kamery, musí být ve vodorovné pozici, aby bylo možno změřit odchylku naklonění hlavy od vodorovné osy. Pokud je kamera vykloněná, není možno přesně vyhodnotit změny polohy hlavy a tím pádem vyhodnotit rizikovou situaci. Metoda používá porovnávání vzájemné polohy dvou vektorů, kde první vektor je tvořen spojnicí mezi určenými body na obličeji a druhý vektor je tvořen vodorovnou polohou. Dle velikosti úhlu, které svírají tyto vektory se určuje stupeň odchylky. Ukázka na obrázcích 29(a) a 29(b).



(a) Naklonění o 1 stupeň.

(b) Naklonění o 32 stupňů.

Obrázek 29: Naklonění hlavy

5.7 Testování

V této části se budeme zabývat testováním přesnosti detekce metod. Každá metoda je otestována na dvou testovacích videích. V prvních videosekvencích je tvář detekovaná z kamery, která snímá tvář z přímého pohledu, ve druhých testovacích videosekvencích snímá kamera tvář v úhlu dvacet stupňů od přímého pohledu z pravé či levé strany. Taký je nutné říct, že všechny testovací videa jsem vytvořil na sobě a na nich vyzkoušel úspěšnost. Testování proběhlo na počítači s procesorem Intel Core I5-7200U, 2.5GHz a 8 GB DDR4 RAM. Videosekvence měly rozlišení 324 x 576 a jeden snímek trval průměrně 60ms.

5.7.1 Asymetrie

Úhel pohledu	Úspěšnost 1.metoda	Úspěšnost 2.metoda(SVM)
0°	98%	92%
20°	90%	73%

Tabulka 1: Výsledky testování nesymetrie tváře obou metod

Z výsledků asymetrie můžeme vidět, že první metoda, která je založena na dělení tváře na identické poloviny podél svislé osy dosahuje vysoké úspěšnosti. Oproti tomu metoda využívající podpůrné vektory nedosahuje tak kvalitních výsledků. Zejména při videu s posunem o 20 stupňů nedetekuje tak moc změny ve tváři. Je to dáno tím, že klasifikátor je natrénovaný datovým setem, který byl vytvořen z frontálního úhlu pohledu. Rychlost výpočtu však mluví opačně. Metoda založená na podpůrných vektorech dokáže získat výsledek velice rychle oproti první metodě. Z mého měření dosahuje větší rychlosti až několikaset násobně.

Úhel pohledu	Falešně pozitivní asymetrie 1.metoda	Falešně pozitivní asymetrie 2.metoda (SVM)
0°	3%	5%
20°	53%	9%

Tabulka 2: Výsledky falešně pozitivních asymetrií

Ve výsledcích tabulky 2 lze vidět, že u normálního běžného výrazu došlo při testování k falešným pozitivním nesymetriím. Vyšší procento těchto chybných vyhodnocení bylo při pohledu z boku, zejména u první metody. Možnou příčinou je, že při bočním pohledu se zájmové body z jedné strany přiblížily k dělicí čáře více, než z opačné strany. Implementovaný algoritmus pro asymetrii se snaží řešit tento problém, avšak jak vidíme, stále se vyskytují falešné pozitivní výsledky. Nadruhou stranu metoda, která využívá podpůrné vektory, má nízkou míru detekce falešně pozitivní asymetrie, ale za cenu nižší detekce pozitivní asymetrie.

5.7.2 Mrkání

K testování detekce mrkání jsem použil testovací video, kdy měření probíhá při stálých světelných podmínkách bez oslnění. Výsledky naznačují vysokou spolehlivost metody, zvláště při přímém pohledu kamery. Výsledky jsou znázorněné v tabulce 3.

Úhel pohledu	Falešně pozitivní mrknutí	Zdetekované mrknutí
0°	4%	93%
20°	2%	89%

Tabulka 3: Výsledky detekce mrkání

5.7.3 Mikrospánek

Při testování frontálního pohledu v případě mikrospánku byly výsledky velmi dobré, úspěšnost dosahovala 94%. Oproti tomu boční pohled vykázal úspěšnost pouze 70%. Při snímání tváře z bočního pohledu došlo několikrát k chybnému vyhodnocení obrazu, kdy došlo k nesprávnému vyhodnocení a záměně mikrospánku za mrkání. Je to způsobeno tím, že algoritmus pracuje na principu hlášení mikrospánku po třech vteřinách zavřených očí. Toto falešné mrkání snižuje úspěšnost detekce spánku.

Úhel pohledu	Úspěšnost
0°	94%
20°	70%

Tabulka 4: Výsledky časové úspěšnosti detekce mikrospánku

5.7.4 Zívání

Při zkoumání zívání bylo zjištěno, že průkaznějších výsledků bylo dosaženo při použití hodnot získaných z čelního pohledu kamery. Při bočním pohledu bylo občas zobrazení zájmových bodů nepřesné, a proto došlo i k nezdetekovatelným a falešně pozitivním závěrům.

Úhel pohledu	Nezdetekované zívnutí	Falešně pozitivní zívnutí
0°	4%	1%
20°	26%	19%

Tabulka 5: Výsledky zívnutí

5.7.5 Třes

V této sekci bylo nezbytné provádět test i pro detekci tváře, protože při třesu se tvář někdy dostane do takové polohy, kde prediktor nedokáže tvář najít. Při provádění testu bylo zjištěno, že při čelním pohledu byly výsledky obecně lepší než při bočním pohledu. Důvodem, proč v druhém případě měření dosáhlo jen 53% úspěšnosti, je to, že z bočního pohledu třes není tak výrazný. V tom případě algoritmu může trvat déle, než dojde ke zjištění třesu. Další nevýhodou je, že při pohledu z boku nejsou zájmové body vždy získatelné, tedy že dochází k výpadkům detekce tváře.

Úhel pohledu	Úspěšnost	Nezdetekované snímky tváře
0°	81%	0%
20°	53%	5.1%

Tabulka 6: Výsledky třesu

5.7.6 Naklonění

V této části testu bylo nezbytné snímky provádět i pro detekci tváře z těch samých důvodů, jako to je u třesu. Při měření změny polohy nakloněním do stran bylo zjištěno, že při stejném úhlu naklonění hlavy se naměřené hodnoty mohou lišit až o 12% v závislosti z jakého pohledu byly údaje získány. Z čelního pohledu údaje vykazovaly vyšší a přesnější hodnoty. Z bočního pohledu bylo vyšší počet nezdetekovaných snímků tváře.

Úhel pohledu	Zdetekovaný úhel	Nezdetekované snímky tváře
0°	40°	7%
20°	35°	18%

Tabulka 7: Výsledky naklonění

5.8 Shrnutí testování

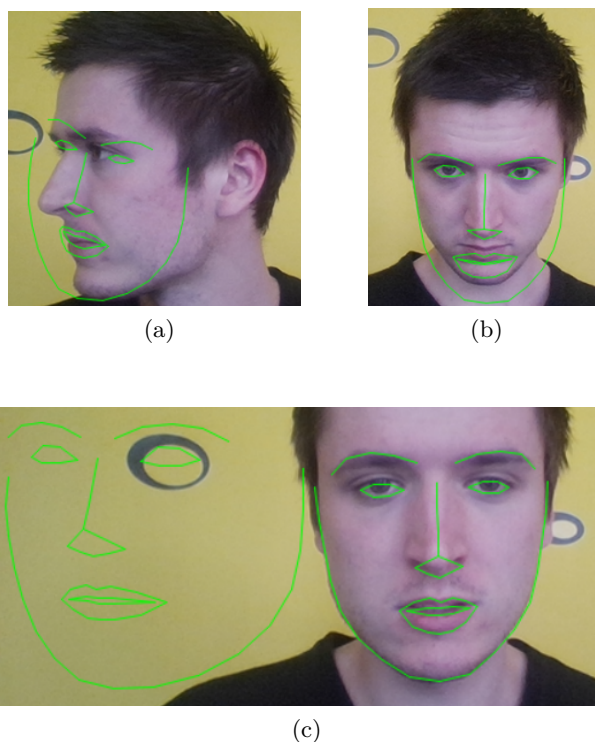
Z tabulkových výsledků jednotlivých akcí můžeme usoudit, že s větším posunem kamery dochází k chybějším rozpoznávání akcí. Důvodů, které k tomu přispívají nebo mohou přispívat hned je několik.

5.8.1 Různý typ tváří a úhel pohledu

Výpočty, které jsou použité, můžou být pro typ lidí s určitou tváří výborné, nadruhou stranu lidé mají širokou škálu tvarů obličejů a tak není možné se dopočítat k dokonalým výsledkům. Dále ke stížení výpočtového algoritmu přispívá natočení hlavy, která výrazně přispívá ke složitosti algoritmů. Čím více řidič pootočí hlavu mimo svou osu, tím více jsou zájmové body blíže sobě, nebo jsou i překryté.

5.8.2 Chybná detekce tváře a zájmových bodů

Nejen samotný úhel pohledu, ale také detekce tváře přispívá k větší chybovosti. Na obrázcích 29(a), 29(b) jdou vidět chybně získané ohraničení tváře. To se stává nejčastěji při natočení hlavy, nebo při detekci se špatným osvětlením. V obrázku 29(c) je zase v reálu znázorněna pouze jedna tvář, ovšem detektor chybně zdetekoval dvě tváře.



Obrázek 29: Chybně určené zájmové body

5.8.3 Příliš malé rozdíly pohledů

Jednotlivé pohledy se můžou velmi hodně podobat. Například řidič, který míjí auto s dálkovými světly podvědomě přiblíží k sobě víčka až o několik desítek procent. To může zaznamenat chybně akci jako mrkání, přitom ale řidič nemrknul. Takových situací je mnoho a je velmi obtížné je správně klasifikovat.

6 Závěr

Úkolem bakalářské práce byla detekce zdravotního stavu řidiče pomocí obrazů. To zahrnovalo seznámení s příznaky zdravotních potíží a taky seznámení s různými metodami pro monitorování a analýzu obličeje a implementace programu pro tuto detekci.

První část této bakalářské práce seznamuje s metodou postavenou na Haarových příznamech od autorů P. Viola a M. Jones, s metodou lokálního binárního vzoru a také s metodou histogramu orientovaných gradientů(HOG). Dále jsou popsány některé metody pro získání zájmových bodů tváře a nakonec je popsáno několik datových setů pro trénování detektorů. V závěru první části je popsán klasifikátor, který je založen na metodě podpůrných vektorů.

Druhá část práce je zaměřená na implementaci a testování. K získání tváře z obrazů byla pomocí knihovny Dlib použita metoda HOG. Pro získání zájmových bodů byla použita metoda představená autory Sullivan a Kazemi, rovněž podporovaná knihovnou Dlib. Byly představeny implementované metody pro zjištění zdravotního stavu řidiče, které následně byly otestovány na testovacích videosekvencích. Výsledky byly vyhodnoceny a na základě výsledků byly popsány výhody a nevýhody implementovaných funkcí.

Pro vylepšení by se dalo využít technologie, které mají větší výpočetní sílu. Pro vylepšení klasifikace asymetrie by se dalo vytvořit datový set z několika desítek nebo stovek snímků lidí, jako trénovací data. Výpočty jsou zaměřené na jediného člověka, takže by se i pro vylepšení mohly změnit algoritmy na analýzu zdravotního stavu více lidí najednou.

Sledování zdravotního stavu řidiče je velice zajímavou oblastí, ve které se dá uplatnit mnoho metod ke zlepšení bezpečnosti řidičů a tím zvýšení bezpečnosti silničního provozu. Tato oblast strojového vidění je každým dnem zlepšovaná, s tím přicházejí přesnější a robustnější detekce algoritmů.

Literatura

- [1] Driver Monitoring: a camera to monitor driver alertness.[Online] Dostupné z:
<https://www.valeo.com/en/driver-monitoring>
- [2] FACIAL RECOGNITION – HISTORY AND PRIVACY.[Online] Dostupné z:
<https://geekforhireinc.com/facial-recognition>
- [3] ZHANG, Fang, et al. Driver fatigue detection based on eye state recognition. In: 2017 International Conference on Machine Vision and Information Technology (CMVIT). IEEE, 2017. p. 105-110.
- [4] OpenCV: Face Detection using Haar Cascades[Online]. OpenCV 3.4.2 Documentation. Dostupné z:
https://docs.opencv.org/3.4.2/d7/d8b/tutorial_py_face_detection.html
- [5] VIOLA, Paul, et al. Rapid object detection using a boosted cascade of simple features. CVPR (1), 2001, 1: 511-518.
- [6] DALAL, Navneet; TRIGGS, Bill. Histograms of oriented gradients for human detection. In: international Conference on computer vision & Pattern Recognition (CVPR'05). IEEE Computer Society, 2005. p. 886–893.
- [7] SHU, Chang; DING, Xiaoqing; FANG, Chi. Histogram of the oriented gradient for face recognition. Tsinghua Science and Technology, 2011, 16.2: 216-224.
- [8] Learn OpenCV: Histogram of Oriented Gradients[Online]. Dostupné z:
<https://www.learnopencv.com/histogram-of-oriented-gradients>
- [9] HUANG, Di, et al. Local binary patterns and its application to facial image analysis: a survey. IEEE Transactions on Systems, Man, and Cybernetics, Part C (Applications and Reviews), 2011, 41.6: 765-781.
- [10] WU, Yue; JI, Qiang. Facial landmark detection: A literature survey. International Journal of Computer Vision, 2019, 127.2: 115-142.
- [11] Facial landmarks with dlib, OpenCV, and Python[Online]. Dostupné z:
<https://www.pyimagesearch.com/2017/04/03/facial-landmarks-dlib-opencv-python/>
- [12] REN, Shaoqing, et al. Face alignment at 3000 fps via regressing local binary features. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 2014. p. 1685-1692.
- [13] KAZEMI, Vahid; SULLIVAN, Josephine. One millisecond face alignment with an ensemble of regression trees. In: Proceedings of the IEEE conference on computer vision and pattern recognition. 2014. p. 1867-1874.

- [14] COOTES, Timothy F., et al. Active shape models-their training and application. Computer vision and image understanding, 1995, 61.1: 38-59.
- [15] IIAU-LAB: Active Shape Model[Online]. Dostupné z:
<http://ice.dlut.edu.cn/lu/ASM.html>
- [16] Intelligent Behaviour Understanding Group (iBUG), Department of Computing, Imperial College London[Online] Dostupné z:
<https://ibug.doc.ic.ac.uk/resources/facial-point-annotations>
- [17] Vuong Le Interactive face segmentation for portrait editing and Helen dataset.[Online] Dostupné z:
<http://www.ifp.illinois.edu/vuongle2/helen>
- [18] 1st Workshop on 3D Face Alignment in the Wild (3DFAW) & Challenge In conjunction with ECCV 2016, Amsterdam, The Netherlands[Online]. Dostupné z:
<http://mhug.disi.unitn.it/workshop/3dfaw>
- [19] "Localizing Parts of Faces Using a Consensus of Exemplars,"Peter N. Belhumeur, David W. Jacobs, David J. Kriegman, Neeraj Kumar, Proceedings of the 24th IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 2011.
- [20] Martin Koestinger, Paul Wohlhart, Peter M. Roth and Horst Bischof In Proc. First IEEE International Workshop on Benchmarking Facial Image Analysis Technologies, 2011.
- [21] Robust face landmark estimation under occlusion X. P. Burgos-Artizzu, P. Perona and P. Dollár ICCV 2013, Sydney, Australia, December 2013.
- [22] Emotion Recognition using Facial Landmarks, Python, DLib and OpenCV by pal-kab[Online]. Převzato z:
<http://www.paulvangent.com/2016/08/05/emotion-recognition-using-facial-landmarks>
- [23] Facial Expression Recognition using Facial Landmark Detection and Feature Extraction via Neural Networks.[Online] Fuzail Khan Department of Electronics and Communication Engineering National Institute of Technology Karnataka. Převzato z:
<https://arxiv.org/pdf/1812.04510.pdf>
- [24] OpenCV: Introduction to Support Vector Machines[Online]. Převzato z:
https://docs.opencv.org/2.4.13.7/doc/tutorials/ml/introduction_to_svm/introduction_to_svm.html
- [25] OpenCV: Support Vector Machines for Non-Linearly Separable Data[Online]. Převzato z:
https://docs.opencv.org/3.1.0/d0/dcc/tutorial_non_linear_svms.html

- [26] OpenCV: About[Online]. Převzato z:
<https://opencv.org/about/>
- [27] Dlib C++ Library: Introduction[Online]. Převzato z:
<http://dlib.net/intro.html>
- [28] Dlib C++ Library: datové soubory a dokumentace[Online]. Převzato z:
<http://dlib.net/files>

A Příloha v IS EDISON

- Zdrojové kódy programu pro detekci zdravotního stavu řidiče
- ReadMe obsahující návod pro spuštění programu